Statistical modeling on $(a, +\infty)$

Pawlowsky-Glahn, V.

Univ. de Girona; Dept. d'Inform. i Matemàtica Aplicada; Campus Montilivi, P4; E-17071 Girona, Spain

Summary. Observations in daily practice are sometimes registered as positive values larger then a given threshold α. The sample space is in this case the interval $(\alpha, +\infty)$, $\alpha > 0$, which can be structured as a real Euclidean space in different ways. This fact opens the door to alternative statistical models depending not only on the assumed distribution function, but also on the metric which is considered as appropriate, *i.e.* the way differences are measured, and thus variability.

1. Introduction and motivation

In this paper I am going to constrain myself to random variables which image or sample space is a generic interval $(\alpha, +\infty)$, $\alpha > 0$. Historically, univariate real statistics has been developed for real random variables, *i.e.* for functions going from a probability space to the real line $\Re$, including those cases where the image set is strictly included in $\Re$. The image or sample space $\Re$ has always been interpreted as the set of real numbers whose elements follow several rules, which are summed up in the definition of Euclidean space. Consequently, it seems natural to sum and multiply elements of $\Re$, to compute the scalar product of two elements, to determine the length or norm of a vector, or to compute the distance between two elements without even thinking if this is the proper way to handle a given data set. Furthermore, there are powerful tools, like integration and derivation, functional relationships, linear algebra, and many more, which allow to solve many problems, at least in an approximate way, although sometimes in a rather complicated manner.

The general theory of linear algebra tells us that any real Hilbert space has an orthonormal basis with respect to which the coefficients behave like usual elements in real space, satisfying all the rules mentioned. Therefrom, it follows that it is enough to proof properties only for the space of coefficients, as they transfer directly to the original space. Hence the usage of the term *Euclidean space* for all those spaces in the finite dimensional case, identifying the properties of real space with the properties of the original space. Of particular interest to us is the fact that concepts like Borel sets, probability and Lebesgue measure, Borel measurable function, probability distribution and density function, moments and moment generating function, and many more, can be taken as defined on the coefficients with respect to an orthonormal basis. In other words, *statistical analysis in an arbitrary Euclidean space can be performed on the coefficients with respect to an orthonormal basis.* This does not mean that the only way to do statistics in such a space is to work on those coefficients, but that they can be used as a shortcut to long mathematical proofs to find properties in any other representation.

One might argue that this facts are of little help in everyday practice, as usual univariate observations *are* real numbers—*i.e.*, they are registered as coefficients of the canonical basis of the real line—and hence *have* to be analyzed using the rules developed for $\Re$. In (Pawlowsky-Glahn et al., in preparation) it is shown that, not denying the fact that usual observations *are* in fact real numbers—*i.e.*, that they are registered as coefficients of a canonical basis of $\Re$—, they *do not have* to be necessarily analyzed applying the rules developed for $\Re$ to the observations themselves. To illustrate the possible practical interest of this approach, there the fact is used that sets like the positive real line, the positive octhant of the real plane, the $(0, 1)$ interval, the unit square, or the sample space of compositional data—data whose measurement units are parts of some whole, like parts per unit, percentages or ppm—can be structured as Euclidean spaces. Consequently, in those very common cases it is possible to apply standard statistical theory to the coefficients in an orthonormal basis. Results differ from standard theory, but are surprisingly easy to obtain and to interpret in some instances, an aspect worthwhile to consider.

Nevertheless, sometimes things do not appear to be that simple, and this is the case presented here. I show that a single set, namely the interval $(\alpha, +\infty) \subseteq \Re_+ \subset \Re$, can be structured in different ways, thus offering different options to statistical modeling. The important point is not the fact that different options are available (there are actually infinitely many), but how to decide which one better suits the data. This is an open question, although my opinion is, that one has to look for the model that best

helps in explaining variability in the data, and this is undissolubly linked to the metric, as shall be seen below.

2. Notation and basic concepts

Consider a one dimensional Euclidean space $\mathcal{E}$. The Abelian group operation or *sum* of two elements $\mathbf{x}, \mathbf{y} \in \mathcal{E}$ will be denoted by $\mathbf{x} \oplus \mathbf{y}$; the iterated *sum* over an index by $\bigoplus_{i=1}^{m} \mathbf{x}_i$; the neutral element with respect to $\oplus$ by ν; the inverse operation, equivalent to subtraction, by $\mathbf{x} \ominus \mathbf{y}$; and the inverse element by $\ominus\mathbf{x}$. Thus, $\mathbf{x} \ominus \mathbf{x} = \nu$. The external multiplication by a scalar $a \in \Re$ will be indicated by $a \odot \mathbf{x}$, and we have $(-1) \odot \mathbf{x} = \ominus\mathbf{x}$, *i.e.* $\mathbf{x} \oplus ((-1) \odot \mathbf{x}) = \mathbf{x} \ominus \mathbf{x} = \nu$. Scalar product, norm and distance will be denoted as usual by $\langle \mathbf{x}, \mathbf{y} \rangle$, $\|\mathbf{x}\|$, $\mathrm{d}(\mathbf{x}, \mathbf{y})$. A subindex will be used only in those cases where needed to avoid confusion. Given that $\mathcal{E}$ is an Euclidean space, if $\mathbf{w}$ is a unitary basis, any observation $\mathbf{x} \in \mathcal{E} \subset \Re$ can be either expressed in terms of its coefficient x in the canonical basis $\mathbf{u}$ of $\Re$, or in terms of its (real) coefficient c_x, in the given basis, *i.e.*

$$\mathbf{x} = x\,\mathbf{u} = c_x \odot \mathbf{w}.$$

From real vector space properties it is known that

$$\mathbf{x} \oplus \mathbf{y} = (c_x + c_y) \odot \mathbf{w}, \quad \text{and} \quad a \odot \mathbf{x} = (a \cdot c_x) \odot \mathbf{w},$$

a fact that suggests the definition of an inner product of two elements, as well as its inverse operation, the inner quotient, as

$$\mathbf{x} \otimes \mathbf{y} = (c_x \times c_x) \odot \mathbf{w}, \quad \text{and} \quad \mathbf{x} \oslash \mathbf{y} = \frac{c_x}{c_y} \odot \mathbf{w},$$

provided that $c_y \neq 0$. They satisfy with respect to the other operations in $\mathcal{E}$ the same properties as the usual inner product and quotient in $\Re$, thus getting a field structure. One important aspect is that it is possible to consider

$$\underbrace{\mathbf{x} \otimes \mathbf{x} \otimes \cdots \otimes \mathbf{x}}_{k \quad \text{times}} = c_x^k \odot \mathbf{w},$$

which is a helpful concept in understanding moments. Obviously, it seems natural to write $\mathbf{x}^k$ for the inner product k times of $\mathbf{x}$ by itself, and to generalize this *power* operation to any real value, thus introducing roots and similar operations in $\mathcal{E}$. Finally, Lebesgue measure in $\mathcal{E}$ associated to a given basis is needed, which is the usual Lebesgue measure on the coefficients. Thus, for $\mathbf{x} = c_x \odot \mathbf{w}$ and $\mathbf{y} = c_y \odot \mathbf{w}$ defining an interval, we get

$$\lambda(\mathbf{x}, \mathbf{y}) = |c_y - c_x|.$$

Given that the basis is fixed, any random variable $\mathbf{X}$ with sample space $\mathcal{E}$ transfers its randomness to its coefficient C_x in that basis. C_x is thus a real random variable in the usual sense, both in the case that $\mathcal{E}$ is coincident with and in the case that it strictly contains the image of $\mathbf{X}$. Furthermore, probability distribution and density functions can be understood as weighting functions, as they multiply the coefficient in a given basis. Consequently, to obtain distributions on $\mathcal{E}$ simply take distributions on those coefficients.

Definition 1. *A* probability distribution function *of a random variable* $\mathbf{X}$ *with sample space* $\mathcal{E}$, $\mathbf{X} = C_x \odot \mathbf{w}$, *is a probability distribution function* $F_{C_x}(c_x)$ *of the random coefficient* C_x *in* $\Re$. *I shall write for short*

$$F_{\mathbf{X}}(\mathbf{x}) = F_{C_x}(c_x),$$

with $\mathbf{x} = c_x\mathbf{w}$. *An analogous definition holds for a probability density function, for which I shall write*

$$f_{\mathbf{X}}(\mathbf{x}) = f_{C_x}(c_x).$$

The probability density function just defined is the Radon-Nykodym derivative of the probability distribution function $F_{\mathbf{X}}(\mathbf{x})$ with respect to the Lebesgue measure in $\mathcal{E}$ at least in the cases studied up to now. I think that this assertion is always true, but it requires the proof that in all cases the Lebesgue measure in $\mathcal{E}$ is absolutely continuous with respect to the usual Lebesgue measure in real space.

It is straightforward, although tedious, to rewrite all the basic properties of functions of a random variables. Therefore, I only recall here some of the most intuitive properties.

PROPOSITION 1. *Let be* $\mathbf{X}$, $\mathbf{Y}$ *random variables with sample space* $\mathcal{E}$ *whose random coefficients with respect to a given unitary basis* $\mathbf{w}$ *are* C_x *and* C_y. *Then it holds*

$$\mathbf{X}\oplus\mathbf{Y}=(C_x+C_y)\odot\mathbf{w};\qquad \mathbf{X}\otimes\mathbf{Y}=(C_x\times C_y)\odot\mathbf{w};$$

$$\mathbf{X}\ominus\mathbf{Y}=(C_x-C_y)\odot\mathbf{w};\qquad \mathbf{X}\oslash\mathbf{Y}=(C_x/C_y)\odot\mathbf{w},\quad \{C_y=0\}=\emptyset;$$

and using standard theory the distribution function or density of any of them can be obtained, whenever the joint density function of $\mathbf{X}$ *and* $\mathbf{Y}$ *is known. Furthermore, for any vector of constants* $\mathbf{a}\in\mathcal{E}$, *any scalar* $a\in\Re$, *any distribution function* $F_{\mathbf{X}}$ *and its corresponding density function* $f_{\mathbf{X}}$, *it holds*

$$F_{\mathbf{a}\oplus\mathbf{X}}(\mathbf{a}\oplus\mathbf{x})=F_{\mathbf{X}}(\mathbf{x});$$

$$F_{a\odot\mathbf{X}}(a\odot\mathbf{x})=F_{\mathbf{X}}(\mathbf{x}),\ \text{if } a>0\,;\quad F_{a\odot\mathbf{X}}(a\odot\mathbf{x})=1-F_{\mathbf{X}}(\mathbf{x}),\ \text{if } a<0\,.$$

$$f_{\mathbf{a}\oplus\mathbf{X}}(\mathbf{a}\oplus\mathbf{x})=f_{\mathbf{X}}(\mathbf{x});\qquad f_{a\odot\mathbf{X}}(a\odot\mathbf{x})=\left|\frac{1}{a}\right|\cdot f_{\mathbf{X}}(\mathbf{x}).$$

Additionally, it is possible to define moments, centred moments and absolute moments of a random variable $\mathbf{X}$ as

$$\mathrm{E}\left[\mathbf{X}^k\right]=\mathrm{E}\left[{\mathbf{C}_\mathbf{x}}^k\right]\odot\mathbf{w};\qquad \mathrm{E}\left[(\mathbf{X}\ominus\mathrm{E}\left[\mathbf{X}\right])^k\right]=\mathrm{E}\left[(\mathbf{C}_\mathbf{x}-\mathrm{E}\left[\mathbf{C}_\mathbf{x}\right])^k\right]\odot\mathbf{w};\mathrm{E}\left[|\mathbf{X}|^k\right]=\mathrm{E}\left[|\mathbf{C}_\mathbf{x}|^k\right]\odot\mathbf{w}$$

Thus, for the mathematical expectation and the second order centred moment, it holds

$$\mathrm{E}\left[\mathbf{X}\right]=\mathrm{E}\left[\mathbf{C}_\mathbf{x}\right]\odot\mathbf{w};\qquad \mathrm{E}\left[(\mathbf{X}\ominus\mathrm{E}\left[\mathbf{X}\right])^2\right]=\mathrm{E}\left[(\mathbf{C}_\mathbf{x}-\mathrm{E}\left[\mathbf{C}_\mathbf{x}\right])^2\right]\odot\mathbf{w}=\mathrm{Var}\left[\mathbf{C}_\mathbf{x}\right]\odot\mathbf{w}.$$

Note that the usual relationship between second order centred moment and first and second non-centred moments holds, as

$$\begin{aligned}\mathrm{E}\left[(\mathbf{X}\ominus\mathrm{E}\left[\mathbf{X}\right])^2\right] &= \mathrm{Var}\left[\mathbf{C}_\mathbf{x}\right]\odot\mathbf{w}=\left(\mathrm{E}\left[{\mathbf{C}_\mathbf{x}}^2\right]-\left(\mathrm{E}\left[\mathbf{C}_\mathbf{x}\right]\right)^2\right)\odot\mathbf{w}\\ &= \left(\mathrm{E}\left[{\mathbf{C}_\mathbf{x}}^2\right]\odot\mathbf{w}\right)\ominus\left(\left(\mathrm{E}\left[\mathbf{C}_\mathbf{x}\right]\right)^2\odot\mathbf{w}\right)\\ &= \mathrm{E}\left[\mathbf{X}^2\right]\ominus\left(\mathrm{E}\left[\mathbf{X}\right]\right)^2.\end{aligned}$$

This fact suggests the definition of measures of variability as the variance, respectively the standard deviation, of the coefficients, whereas moments, both centred and non-centred or functions thereof, are points in space which satisfy certain conditions. They are coincident only in case $\mathcal{E}=\Re$. Thus, $\mathrm{E}\left[(\mathbf{X}\ominus\mathrm{E}\left[\mathbf{X}\right])^2\right]$ is an elements of $\mathcal{E}$ which is the squared expected distance apart from the origin, whereas its square root will be one standard deviation apart from it. The *coefficient of variation* can now be defined as a quotient of coefficients (that of the square root of the second order centred moment and that of the expected value) or as a coefficient of a quotient of elements: in any case it will be a coefficient and the value will be the same. A similar argument holds for measures of skewness, curtosis, and other measures, but I am not going to pursue these aspects further, as they are not the central issue of this paper.

As can be seen, properties transfer easily from the space of coefficients to $\mathcal{E}$ when dealing with probability distribution and density functions, at least in the univariate case. Nevertheless, the essential thing for my purposes is, at this stage, that the study of those functions is mainly the study of some numerical characteristics or parameters associated with them, and I will show how to do that easily in the different possible structures of $(\alpha,+\infty)$.

3. Examples

In this section, three possible Euclidean space structures for $(\alpha,+\infty)\subseteq\Re_+\subset\Re$ are introduced. Proofs are omitted, as they are straightforward, although sometimes tedious. To simplify the presentation, I will use for the logarithm base a, $a\in\Re_+$, the notation lg_a and for its inverse function $\exp_a$. Furthermore, for iterated logarithms, respectively exponentials, I will use an exponent. Thus, for any real value x,

$$\mathrm{lg}_a(x)=\log_a(x);\qquad \mathrm{lg}_a^2(x)=\mathrm{lg}_a(\mathrm{lg}_a(x));\qquad \mathrm{lg}_a^3=\mathrm{lg}_a\left(\mathrm{lg}_a(\mathrm{lg}_a(x))\right);$$

$$\exp_a(x)=a^x;\qquad \exp_a^2(x)=\exp_a\left(\exp_a(x)\right)=a^{a^x};\qquad \exp_a^3(x)=\exp_a\left(\exp_a(\exp_a(x))\right)=a^{a^{a^x}}.$$

Case 1. Recall that every element $\mathbf{x} \in (\alpha, +\infty)$ can be viewed as a real vector, $\mathbf{x} = x\,\mathbf{u}$, $x > 0$, where $\mathbf{u}$ stands for the unit vector in the real line and x for the coefficient in this basis or, as shall be seen, as an element of the Euclidean space $(\alpha, +\infty)$. The example presented in the first place is directly related to the usual transformation used for the three-parameter lognormal distribution.

- Abelian group operation $\oplus$, neutral element ν and inverse element $\ominus\mathbf{x}$:

$$\mathbf{x} \oplus \mathbf{y} = (\alpha + (x-\alpha)(y-\alpha))\ \mathbf{u}; \quad \nu = (\alpha+1)\,\mathbf{u}; \quad \ominus\mathbf{x} = \left(\alpha + \frac{1}{x-\alpha}\right)\ \mathbf{u}.$$

- External multiplication $\odot$:

$$a \odot \mathbf{x} = (\alpha + (x-\alpha)^a)\ \mathbf{u}.$$

- Scalar product $\langle\cdot,\cdot\rangle$, norm $\|\cdot\|$ and distance $\mathrm{d}(\cdot,\cdot)$:

$$\langle \mathbf{x}, \mathbf{y}\rangle = \lg_e(x-\alpha)\lg_e(y-\alpha); \quad \|\mathbf{x}\| = |\lg_e(x-\alpha)|;$$

$$\mathrm{d}(\mathbf{x},\mathbf{y}) = |\lg_e(y-\alpha) - \lg_e(x-\alpha)| = \left|\lg_e \frac{y-\alpha}{x-\alpha}\right|.$$

- Unitary basis $\mathbf{w}$ and coefficient c_x of an arbitrary vector $\mathbf{x}$ in the given basis:

$$\mathbf{w} = (\alpha + e)\,\mathbf{u}\,; \quad c_x = \lg_e(x-\alpha),$$

i.e.

$$\mathbf{x} = c_x \odot \mathbf{w} = \lg_e(x-\alpha) \odot ((\alpha+e)\,\mathbf{u}) = \left(\alpha + e^{\lg_e(x-\alpha)}\right)\ \mathbf{u} = x\,\mathbf{u}.$$

- Internal multiplication $\otimes$ and quotient $\oslash$:

$$\mathbf{x} \otimes \mathbf{y} = (c_x c_y) \odot \mathbf{w} = (\lg_e(x-\alpha)\ \lg_e(y-\alpha)) \odot ((\alpha+e)\,\mathbf{u}) = \alpha + \exp_e(\lg_e(x-\alpha)\ \lg_e(y-\alpha))\ \mathbf{u};$$

$$\mathbf{x} \oslash \mathbf{y} = \frac{c_x}{c_y} \odot \mathbf{w} = \left(\frac{\lg_e(x-\alpha)}{\lg_e(y-\alpha)}\right) \odot ((\alpha+e)\,\mathbf{u}) = \alpha + \exp_e\left(\frac{\lg_e(x-\alpha)}{\lg_e(y-\alpha)}\right)\ \mathbf{u};$$

- Lebesgue measure of an interval defined by the origin ν and an arbitrary point $\mathbf{x}$:

$$\lambda = \lambda(\nu, \mathbf{x}) = |c_x| = |\lg_e(x-\alpha)|.$$

Note that we obtain an analogous structure using any base $a \in \mathbb{R}_+$ for the logarithm and its inverse function. Note also that it has no influence on the vector space structure, but that it determines certainly the scalar product, the norm and the distance, and thus the unitary basis and the corresponding coefficients.

Case 2. Again, consider every element $\mathbf{x} \in (\alpha, +\infty)$ as a real vector, $\mathbf{x} = x\,\mathbf{u}$, $x > 0$, where $\mathbf{u}$ stands for the unit vector in the real line and x for the coefficient in this basis or as an element of the Euclidean space $(\alpha, +\infty)$. This second example has been suggested by the following chain of transformations:

$$\begin{array}{ccccccc} (\alpha,+\infty) & \longrightarrow & (1,+\infty) & \longrightarrow & (0,+\infty) & \longrightarrow & \Re \\ \mathbf{x} = x\,\mathbf{u} & \mapsto & (x/\alpha)\,\mathbf{u} & \mapsto & \lg_e(x/\alpha)\,\mathbf{u} & \mapsto & \lg_e^2(x/\alpha)\,\mathbf{u}. \end{array}$$

The inverse chain then reads as follows:

$$\begin{array}{ccccccc} \Re & \longrightarrow & (0,+\infty) & \longrightarrow & (1,+\infty) & \longrightarrow & (\alpha,+\infty) \\ \mathbf{y} = y\,\mathbf{u} & \mapsto & \exp_e(y)\,\mathbf{u} & \mapsto & \exp_e^2(y)\,\mathbf{u} & \mapsto & \alpha\exp_e^2(y)\,\mathbf{u} = \alpha e^{e^y}\,\mathbf{u}. \end{array}$$

Note that I have used the natural logarithm and its inverse, the exponential base e function, although any other base could be used both for the logarithm and its inverse function.

- Abelian group operation $\oplus$, neutral element ν and inverse element $\ominus\mathbf{x}$:

$$\mathbf{x} \oplus \mathbf{y} = \left(\alpha \exp_e\left(\lg_e \frac{x}{\alpha}\lg_e \frac{y}{\alpha}\right)\right)\ \mathbf{u}; \quad \nu = (ae)\,\mathbf{u}; \quad \ominus\mathbf{x} = \left(\alpha \exp_e\left(\frac{1}{\lg_e(x/\alpha)}\right)\right)\ \mathbf{u}.$$

- External multiplication $\odot$:
$$a\odot\mathbf{x}=\left(\alpha\exp_e\left(\lg_e\frac{x}{\alpha}\right)^a\right)\mathbf{u}.$$

- Scalar product $\langle\cdot,\cdot\rangle$, norm $\|\cdot\|$ and distance $\mathrm{d}(\cdot,\cdot)$:
$$\langle\mathbf{x},\mathbf{y}\rangle=\lg_e^2\frac{x}{\alpha}\lg_e^2\frac{y}{\alpha};\quad \|\mathbf{x}\|=\left|\lg_e^2\frac{x}{\alpha}\right|;$$
$$\mathrm{d}(\mathbf{x},\mathbf{y})=|\lg_e^2\frac{y}{\alpha}-\lg_e^2\frac{x}{\alpha}|=\left|\lg_e\frac{\lg_e\frac{y}{\alpha}}{\lg_e\frac{x}{\alpha}}\right|=\left|\lg_e\frac{\lg_e y-\lg_e\alpha)}{\lg_e x-\lg_e\alpha)}\right|.$$

- Unitary basis $\mathbf{w}$ and coefficient c_x of an arbitrary vector $\mathbf{x}$ in the given basis:
$$\mathbf{w}=\alpha e^e\,\mathbf{u}\,;\quad c_x=\lg_e^2\frac{x}{\alpha},$$
i.e.
$$\mathbf{x}=c_x\odot\mathbf{w}=\lg_e^2\frac{x}{\alpha}\odot(ae^e\,\mathbf{u})=\left(\alpha\exp_e^2\left(\lg_e^2\frac{x}{\alpha}\right)\right)\mathbf{u}=\left(\alpha e^{e^{\lg_e\lg_e\frac{x}{\alpha}}}\right)\mathbf{u}=x\,\mathbf{u}.$$

- Internal multiplication $\otimes$ and quotient $\oslash$:
$$\mathbf{x}\otimes\mathbf{y}=(c_xc_y)\odot\mathbf{w}=\left(\lg_e^2\frac{x}{\alpha}\,\lg_e^2\frac{y}{\alpha}\right)\odot(ae^e\,\mathbf{u})\,;$$
$$\mathbf{x}\oslash\mathbf{y}=\frac{c_x}{c_y}\odot\mathbf{w}=\left(\lg_e^2\frac{x}{\alpha}/\lg_e^2\frac{y}{\alpha}\right)\odot(ae^e\,\mathbf{u})\,;$$

- Lebesgue measure of an interval defined by the origin ν and an arbitrary point $\mathbf{x}$:
$$\lambda=\lambda(\nu,\mathbf{x})=|c_x|=\left|\lg_e^2\frac{x}{\alpha}\right|.$$

Case 3. As before, consider every element $\mathbf{x}\in(\alpha,+\infty)$ as a real vector, $\mathbf{x}=x\,\mathbf{u}$, $x>0$, where $\mathbf{u}$ stands for the unit vector in the real line and x for the coefficient in this basis or as an element of the Euclidean space $(\alpha,+\infty)$. This third example is based on the following chain of transformations:

$$\begin{array}{ccccccc}(\alpha,+\infty) & \longrightarrow & (1,+\infty) & \longrightarrow & (0,+\infty) & \longrightarrow & \Re\\ \mathbf{x}=x\,\mathbf{u} & \mapsto & \lg_\alpha x\,\mathbf{u} & \mapsto & \lg_\alpha^2 x\,\mathbf{u} & \mapsto & \lg_\alpha^3 x\,\mathbf{u}.\end{array}$$

The inverse chain reads as follows:

$$\begin{array}{ccccccc}\Re & \longrightarrow & (0,+\infty) & \longrightarrow & (1,+\infty) & \longrightarrow & (\alpha,+\infty)\\ \mathbf{y}=y\,\mathbf{u} & \mapsto & \exp_\alpha y\,\mathbf{u} & \mapsto & \exp_\alpha^2 y\,\mathbf{u} & \mapsto & \exp_\alpha^3 y\,\mathbf{u}.\end{array}$$

In the following developments I use the same base α for all cases, although looking at the chain above it is clear that only the first step, respectively the last one in the second case, requires the base to be α.

- Abelian group operation $\oplus$, neutral element ν and inverse element $\ominus\mathbf{x}$:
$$\mathbf{x}\oplus\mathbf{y}=\left(\exp_\alpha^2\left(\lg_\alpha^2 x\lg_\alpha^2 y\right)\right)\mathbf{u};\quad \nu=\alpha^\alpha\,\mathbf{u};\quad \ominus\mathbf{x}=\left(\exp_\alpha^2\left(\lg_\alpha^2\frac{1}{x}\right)\mathbf{u}\right).$$

- External multiplication $\odot$:
$$a\odot\mathbf{x}=\left(\exp_\alpha^3\left(a\,\lg_\alpha^3 x\right)\right)=\left(\exp_\alpha^2\left(\lg_\alpha^2 x\right)^a\right)\mathbf{u}.$$

- Scalar product $\langle\cdot,\cdot\rangle$, norm $\|\cdot\|$ and distance $\mathrm{d}(\cdot,\cdot)$:
$$\langle\mathbf{x},\mathbf{y}\rangle=\lg_\alpha^3 x\lg_\alpha^3 y;\quad \|\mathbf{x}\|=|\lg_\alpha^3 x|;$$
$$\mathrm{d}(\mathbf{x},\mathbf{y})=|\lg_\alpha^3 y-\lg_\alpha^3 x|=\left|\lg_\alpha\frac{\lg_\alpha^2 y}{\lg_\alpha^2 x}\right|.$$

- Unitary basis $\mathbf{w}$ and coefficient c_x of an arbitrary vector $\mathbf{x}$ in the given basis:

$$\mathbf{w} = \left(\exp_\alpha^3 1\right)\,\mathbf{u} = \left(\exp_\alpha^2 \alpha\right)\,\mathbf{u} = \left(\alpha^{\alpha^\alpha}\right)\,\mathbf{u}\,; \qquad c_x = \lg_\alpha^3 x,$$

i.e.

$$\mathbf{x} = c_x \odot \mathbf{w} = \lg_\alpha^3 x \odot \left(\left(\exp_\alpha^3 1\right)\,\mathbf{u}\right) = \left(\exp_\alpha^3 \left(\lg_\alpha^3 x\right)\right)\,\mathbf{u} = x\,\mathbf{u}.$$

- Internal multiplication $\otimes$ and quotient $\oslash$:

$$\mathbf{x} \otimes \mathbf{y} = \left(\lg_\alpha^3 x\ \lg_\alpha^3 y\right) \odot \left(\left(\exp_\alpha^3 1\right)\,\mathbf{u}\right) = \left(\exp_\alpha^3 \left(\lg_\alpha^3 x\ \lg_\alpha^3 y\right)\right)\,\mathbf{u};$$

$$\mathbf{x} \oslash \mathbf{y} = \left(\frac{\lg_\alpha^3 x}{\lg_\alpha^3 y}\right) \odot \left(\left(\exp_\alpha^3 1\right)\,\mathbf{u}\right) = \left(\exp_\alpha^3 \left(\frac{\lg_\alpha^3 x}{\lg_\alpha^3 y}\right)\right)\,\mathbf{u}.$$

- Lebesgue measure of an interval defined by the origin ν and an arbitrary point $\mathbf{x}$:

$$\lambda = \lambda\left(\nu, \mathbf{x}\right) = |c_x| = |\lg_\alpha^3 x|.$$

4. The normal distribution on $(\alpha, +\infty)$.

4.1. Density and moments

Let us assume a normal probability distribution function on $(\alpha, +\infty)$ for a random variable $\mathbf{X} = C \odot \mathbf{w}$ or, equivalently, on $\mathbb{R}$ for the coefficient C. Thus, at each $\mathbf{x} \in (\alpha, +\infty)$ we can write $\mathbf{x} = c\,\mathbf{w}$, and the value of the density function at that point will be

$$f_{\mathbf{X}}(\mathbf{x}) = f_C(c) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(c-\mu)^2}{2\sigma^2}\right\},$$

where μ and σ are the parameters of the distribution, which depends also on the value of α through c. Thus, we have a three-parameter normal model, which we shall denote as $\mathcal{N}_\alpha(\mu, \sigma)$. We know that for this model moments about the mean, about the origin, and absolute moments of all order exist. They can be easily obtained from the moments of the coefficients, which are those of a normal random variable. For instance, if k is even

$$\mathrm{E}\left[\left(\mathbf{X} - \mathrm{E}\left[\mathbf{X}\right]\right)^k\right] = \mathrm{E}\left[\left(C - \mathrm{E}\left[C\right]\right)^k\right] \odot \mathbf{w} = \left((k-1)\cdot(k-2)\cdots 3\cdot 1\cdot\sigma^k\right) \odot \mathbf{w},$$

but if k is odd

$$\mathrm{E}\left[\left(\mathbf{X} - \mathrm{E}\left[\mathbf{X}\right]\right)^k\right] = 0 \odot \mathbf{w}.$$

Furthermore, the mean, the median and the mode are coincident and are equal to $\mu \odot \mathbf{w}$, and the distribution is symmetric around them. This is not easy to appreciate from an histogram, but one has to be aware that we are working with metrics different from the usual Euclidean metric in $\mathbb{R}$.

4.2. Parameter estimation

The main problem in using this, and any other distribution, is to obtain a reasonable estimate of the three parameters involved. Many different methods can be found in the literature to solve this problem, at least by approximation. Here I am going to follow the scheme in **?**), and in each part I will present first the general results, which are independent of the space structure assumed, and then the details for each of the three cases considered.

4.2.1. Maximum likelihood function

General. If we consider the loglikelihood function for a sample, we obtain easily the usual ML estimates m_ℓ for μ and s_ℓ^2 for σ^2 expressed in terms of coefficients:

$$m_\ell = \frac{1}{n}\sum_{i=1}^n c_i\,; \tag{1}$$

$$s_\ell^2 = \frac{1}{n}\sum_{i=1}^n c_i^2 - m_\ell^2. \tag{2}$$

The parameter α is included in the coefficients c_i and, therefore, there is no general solution a_ℓ for α. The equation to be solved is

$$\sum_{i=1}^{n} \left(c_i - \frac{1}{n} \sum_{i=1}^{n} c_i \right) \frac{\partial c_i}{\partial \alpha} = 0. \tag{3}$$

From here on it is necessary to analyze the three cases separately, and I will concentrate on the function we need to solve to obtain a_ℓ.

Case 1. We have $c = \lg_e(x - \alpha)$ and thus $\partial c/\partial \alpha = -(x - \alpha)^{-1}$ leading to

$$\sum_{i=1}^{n} \frac{\lg_e(x_i - \alpha) - \frac{1}{n} \sum_{j=1}^{n} \lg_e(x_j - \alpha)}{x_i - \alpha} = 0.$$

Case 2. We have $c = \lg_e^2(x/\alpha)$ and thus $\partial c/\partial \alpha = -\left(\alpha(\lg_e x - \lg_e \alpha)\right)^{-1}$ leading to

$$\sum_{i=1}^{n} \frac{\lg_e^2(x_i/\alpha) - \frac{1}{n} \sum_{i=1}^{n} \lg_e^2(x_i/\alpha)}{\lg_e x_i - \lg_e \alpha} = 0.$$

Case 3. We have $c = \lg_\alpha^3(x)$ and thus $\partial c/\partial \alpha = -(x \lg_\alpha x \lg_\alpha^2 x)^{-1}$ leading to

$$\sum_{i=1}^{n} \frac{\lg_\alpha^3(x_i) - \frac{1}{n} \sum_{i=1}^{n} \lg_\alpha^3(x_i)}{x \lg_\alpha x \lg_\alpha^2 x} = 0.$$

Note that the first two equations have, with respect to the coefficients, the same form. In fact, they could both be written

$$\sum_{i=1}^{n} \frac{c_i - \bar{c}}{\exp_e c_i} = 0,$$

whereas for the third one we have

$$\sum_{i=1}^{n} \frac{c_i - \bar{c}}{\exp_\alpha c_i \cdot \exp_\alpha^2 c_i \cdot \exp_\alpha^3 c_i} = 0.$$

4.2.2. Cohen's least sample method

General. It implies keeping the general solution equations (1) and (2) and substituting (3) by one based on the least sample value, x_0, which is assumed to have empirical frequency n_0 in n samples. Thus, the general estimators are obtained from

$$m_c = \frac{1}{n} \sum_{i=1}^{n} c_i = \bar{c}; \tag{4}$$

$$s_c^2 = \frac{1}{n} \sum_{i=1}^{n} c_i^2 - m_2^2; \tag{5}$$

$$x_0 = c_0 \odot \mathbf{w} = (m_2 + \nu s_2) \odot \mathbf{w} \iff c_0 - m_2 - \delta s_2 = 0, \tag{6}$$

where δ is the quantile n_0/n from the standard normal. The estimate α_c of α will thus be the solution to the general equation

$$c_0 - \frac{1}{n} \sum_{i=1}^{n} c_i - \frac{\delta}{n} \sum_{i=1}^{n} c_i^2 - \left(\frac{1}{n} \sum_{j=1}^{n} c_j \right)^2 = 0,$$

which is written in each of the three cases considered as follows.

Case 1. $c = \lg_e(x - \alpha)$:

$$\lg_e(x_0 - \alpha) - \frac{1}{n} \sum_{i=1}^{n} \lg_e(x_i - \alpha) - \frac{\delta}{n} \sum_{i=1}^{n} \left(\lg_e(x_i - \alpha)\right)^2 - \left(\frac{1}{n} \sum_{j=1}^{n} \lg_e(x_j - \alpha) \right)^2 = 0.$$

Case 2. $c = \lg_e^2(x/\alpha)$:

$$\lg_e^2(x_0/\alpha) - \frac{1}{n}\sum_{i=1}^{n}\lg_e^2(x_i/\alpha) - \frac{\delta}{n}\sum_{i=1}^{n}\left(\lg_e^2(x_i/\alpha)\right)^2 - \left(\frac{1}{n}\sum_{j=1}^{n}\lg_e^2(x_j/\alpha)\right)^2 = 0.$$

Case 3. $c = \lg_\alpha^3(x)$:

$$\lg_\alpha^3(x_0) - \frac{1}{n}\sum_{i=1}^{n}\lg_\alpha^3(x_i) - \frac{\delta}{n}\sum_{i=1}^{n}\left(\lg_\alpha^3(x_i)\right)^2 - \left(\frac{1}{n}\sum_{j=1}^{n}\lg_\alpha^3(x_j)\right)^2 = 0.$$

4.2.3. Method of moments

General. The estimators for μ and σ^2 are obtained from the coefficients as the usual arithmetic means, one of observations, the other of squared deviations from the mean, whereas the estimator for the parameter α is obtained including the third centred moment, which theoretical value is zero. Thus, the general estimators are obtained from

$$m_m = \frac{1}{n}\sum_{i=1}^{n} c_i = \bar{c}; \tag{7}$$

$$s_m^2 = \frac{1}{n}\sum_{i=1}^{n} c_i^2 - \left(\frac{1}{n}\sum_{j=1}^{n} c_j\right)^2; \tag{8}$$

$$0 = \frac{1}{n}\sum_{i=1}^{n}(c_i - m_m)^3. \tag{9}$$

For the three cases considered we obtain thus as an estimate of α the solution α_m of the following equations

Case 1. $c = \lg_e(x-\alpha)$:

$$\frac{1}{n}\sum_{i=1}^{n}(\lg_e(x_i-\alpha) - \frac{1}{n}\sum_{i=1}^{n}\lg_e(x_i-\alpha))^3 = 0.$$

Case 2. $c = \lg_e^2(x/\alpha)$:

$$\frac{1}{n}\sum_{i=1}^{n}\lg_e^2(x_i/\alpha) - \frac{1}{n}\sum_{i=1}^{n}\lg_e^2(x_i/\alpha))^3 = 0.$$

Case 3. $c = \lg_\alpha^3(x_i)$:

$$\frac{1}{n}\sum_{i=1}^{n}\lg_\alpha^3(x_i) - \frac{1}{n}\sum_{i=1}^{n}\lg_\alpha^3(x_i)))^3 = 0.$$

4.2.4. Method of quantiles

General. In this method, all three required estimates are obtained directly from the sample quantiles, which are directly obtained from the observations. Thus, if we look for the quantiles q, 0.5 and $1-q$, we have the general equations:

$$\mathbf{x}_{0.5} = c_{0.5}\cdot\mathbf{w} = m_q\cdot\mathbf{w}; \tag{10}$$

$$\mathbf{x}_q = c_q\cdot\mathbf{w} = (m_q - \delta s_q)\cdot\mathbf{w}; \tag{11}$$

$$\mathbf{x}_{1-q} = c_{1-q}\cdot\mathbf{w} = (m_q + \delta s_q)\cdot\mathbf{w}; \tag{12}$$

where $\delta \in [0, 1]$ stands for the theoretical value of the q-th quantile of the standard normal distribution. Considering only the coefficients we have

$$\begin{aligned} m_q &= c_{0.5}; && (13)\\ s_q &= \frac{c_{1-q} - c_q}{2\delta}; && (14)\\ 0 &= c_{1-q} + c_q - 2c_{0.5}, && (15) \end{aligned}$$

and with this approach s_q will always be positive by construction. For the three cases considered we obtain the following estimators:

Case 1. $c = \lg_e(x - \alpha)$:

$$\begin{aligned} m_q &= \lg_e(x_{0.5} - \alpha_q); && (16)\\ s_q &= \frac{\lg_e(x_{1-q} - \alpha_q) - \lg_e(x_q - \alpha_q)}{2\delta}; && (17)\\ \alpha_q &= \frac{x_{1-q}x_q - x_{0.5}^2}{x_{1-q} + x_q - 2x_{0.5}}. && (18) \end{aligned}$$

Case 2. $c = \lg_e^2(x/\alpha)$:

$$\begin{aligned} m_q &= \lg_e^2(x_{0.5}/\alpha_q); && (19)\\ s_q &= \frac{\lg_e^2(x_{1-q}/\alpha_q) - \lg_e^2(x_q/\alpha_q)}{2\delta}; && (20)\\ a_q &= \exp_e\left(\frac{\lg_e x_{1-q} \lg_e x_q - (\lg_e x_{0.5})^2}{\lg_e x_{1-q} + \lg_e x_q - 2\lg_e x_{0.5}}\right). && (21) \end{aligned}$$

Case 3. $c = \lg_\alpha^3(x_i)$:

$$\begin{aligned} m_q &= \lg_{\alpha_q}^3(x_{0.5}); && (22)\\ s_q &= \frac{\lg_{\alpha_q}^3(x_{1-q}) - \lg_{\alpha_q}^3(x_q)}{2\delta}; && (23)\\ \alpha_q &= \exp_e^2\left(\frac{\lg_e^2 x_{1-q} \lg_e^2 x_q - (\lg_e^2 x_{0.5})^2}{\lg_e^2 x_{1-q} + \lg_e^2 x_q - 2\lg_e^2 x_{0.5}}\right). && (24) \end{aligned}$$

References